

Comparison of the heavy-fermion and Foldy-Wouthuysen formalisms at third order

A. Gårdestig,^{*} K. Kubodera,[†] and F. Myhrer[‡]

*Department of Physics and Astronomy,
University of South Carolina, Columbia, SC 29208*

(Dated: February 5, 2008)

Abstract

We compare two non-relativistic (NR) reduction schemes (heavy-fermion and Foldy-Wouthuysen) that are used to derive low-energy effective-field-theory Lagrangians. We give the explicit transformation between the two types of fields to $\mathcal{O}(1/m^2)$, derived from a quite general, relativistic Lagrangian. Beyond leading order the NR reductions always involve the smaller components of the Dirac spinors that are to be integrated out to formulate the NR theory. Even so, the transformation between the NR Lagrangians can be carried out explicitly to $\mathcal{O}(1/m^2)$ using a field renormalization, as long as the lower components of the Lagrangian are known. The fixed coefficient corrections to some low-energy constants at $\mathcal{O}(1/m^2)$ will depend on the particular scheme chosen, but will match after the field renormalization.

PACS numbers: 11.10.-z, 11.25.Db

Keywords: effective field theory, non-relativistic reductions, Foldy-Wouthuysen transformation, heavy-fermion method

^{*}Electronic address: anders@physics.sc.edu

[†]Electronic address: kubodera@sc.edu

[‡]Electronic address: myhrer@physics.sc.edu

I. INTRODUCTION

When considering low-energy phenomena, it often simplifies the calculations to start from a non-relativistic Lagrangian. Several useful and well-known examples can be found. These include the low-energy few-body nuclear systems for which chiral perturbation theory (χ PT) has been developed, the electromagnetic properties of composite particles (moments of the proton and neutron), atomic physics [non-relativistic quantum electrodynamics (NRQED)] and heavy-quark effective field theory (HQEFT). Since the original theory, here assumed to be an effective field theory (EFT) in the Weinberg sense [1], is generally given by a relativistic Lagrangian, one faces the problem of performing a non-relativistic (NR) reduction.

Several solutions have been suggested over the years. The simplest, which we will label the direct Pauli reduction scheme or the $1/m$ expansion, expands the Dirac matrix element in terms of the two-dimensional Pauli-spinors that appear in the usual representation of the four-spinors.

HQEFT was originally formulated in the non-relativistic limit, utilizing the constraints given by the symmetries of quantum chromodynamics (QCD) in the static limit [2]. Later HQEFT was elegantly derived from the relativistic QCD Lagrangian using a path-integral method [3]; this we will call the heavy-fermion (HF) approach (for a recent review of HQEFT, see [4]). These ideas were carried over to the descriptions of heavy baryon chiral perturbation theory (HB χ PT) [5, 6, 7].

An alternative NR reduction method would be the classic Foldy-Wouthuysen (FW) transformation [8, 9], which was designed to provide the NR limit of quantum electrodynamics (QED). It has also been used to derive HQEFT [10], but this work was initially criticized by Ref. [3] for not proving HQEFT to be an EFT (see, however, later work [11, 12]). According to Ref. [3], HF and FW give identical results for HQEFT, although as far as we are able to ascertain, this was shown only to $\mathcal{O}(1/m)$, where m is the (heavy) fermion mass. Later, these two methods were indeed shown to be equivalent, after a wave function renormalization and a unitary transformation [13, 14] (see also Refs. [15, 16]). However, these references considered only renormalizable theories (QED and QCD) and [14] only the explicit example of a constant electric field (in QED). Neither considered the low-energy constants (LECs), which parametrize the (usually unknown) short-distance physics in an EFT like, e.g., χ PT.

It has already been shown that, when the relativistic theory is fully known, the direct

Pauli reduction scheme ($1/m$ expansion) differs from FW and only the latter gives S -matrices that are in agreement with the relativistic ones [17]. This provided part of the inspiration for the present investigation. Another motivation was the claims in the literature that FW and HF are equivalent for HQEFT [3, 10] as well as for HB χ PT [18].

In both HF and FW the contributions involving the lower Dirac spinor components are eliminated, and this can be formalized by integrating out these “small” components from the path integral. For momenta and field energies smaller than m , the surviving terms can be expanded in a perturbative series. The difference between the two methods lies in the way the separation of lower and upper components is made, resulting in different NR fields. Note that this ‘integrating-out’ procedure is different from what is done in, e.g., electroweak theory, where the heavy vector bosons are integrated out completely, at lowest order giving the Fermi weak interaction model. Here we rather want to keep the heavy (fermion) degrees of freedom (when they are protected by a symmetry, like baryon number conservation), but eliminate the explicit antifermion contributions (lower components), since they can not be excited in the non-relativistic limit.

In this paper we will study the HF and FW representations in detail, starting from a quite general, relativistic, Lagrangian, which is assumed to be known. The two methods appear to disagree at higher orders in $1/m$, i.e., the so-called “fixed-coefficient” terms ($1/m$ corrections) differ between the Lagrangians. We will give explicit expressions for the transformation necessary to get from HF to FW to $\mathcal{O}(1/m^2)$, showing that the corresponding fields each contain pieces of the small/lower components of the other, already at $\mathcal{O}(1/m)$. Even so, the functional form of the corresponding Lagrangians will differ first at $\mathcal{O}(1/m^2)$, which explains the equivalence found between the $\mathcal{O}(1/m)$ results of Refs. [3] and [10]. It is possible to show the equivalence of the two approaches [14] (at least for a constant electric field in NRQED), by using field renormalization and unitary transformations. We want to treat a more general Lagrangian than what was done in [14] and will show that at higher orders there are differences in the fixed-coefficient ($1/m$) corrections to the LECs in, e.g., χ PT. (In this paper we will only be concerned with LECs introduced already in the relativistic Lagrangian, as done in, e.g., [7, 19, 20].) Thus, one has to be careful to use correctly renormalized fields when comparing LECs at this order. At higher order also unitary transformations need to be applied.

This paper is organized as follows: In Sec. II we give a brief review of the essential

features of the HF and FW schemes, followed by a detailed comparison and derivation of the transformation between them. The paper is concluded by a discussion in Sec. III. Some mathematical details are collected in an Appendix.

II. NON-RELATIVISTIC REDUCTIONS

We will assume as the starting point the relativistic Lagrangian

$$\mathcal{L} = \bar{\psi} (i\mathcal{D} - m + \gamma_0 G) \psi, \quad (1)$$

where the covariant derivative $\mathcal{D}^\mu \equiv \partial^\mu + \Gamma^\mu$ and the general coupling matrix G (the extra factor of γ_0 is included for future convenience) is

$$G \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (2)$$

where A – D are 2×2 matrices. The connection Γ^μ and the coupling matrix G contain all single-fermion couplings (to, e.g., pions and electroweak fields) allowed by relevant symmetries, e.g., chiral symmetry, Lorentz invariance, parity, flavor and color SU(3) etc. Throughout the paper we will work in the usual Dirac representation [9] and G should be interpreted in that representation. We further assume that G is hermitian, i.e., that A and D are hermitian and $B^\dagger = C$. No assumptions will be made regarding the commutative properties of the elements of G and \mathcal{D} , so the results of this paper apply equally well to Abelian and non-Abelian field theories. Note that A – D and \mathcal{D}^μ will in general contain higher order interactions terms which should be expanded in the final NR formulas. This expansion is separate from the $1/m$ expansion created by the NR reduction, but in some cases the two expansions can be combined into a single counting scheme. This happens for, e.g., HB χ PT, where $m \sim \Lambda_\chi$ and $\Lambda_\chi \sim 1$ GeV is the chiral scale. The generating functional for the Lagrangian (1) is

$$\mathcal{Z} = \int [d\bar{\psi}][d\psi][dX] \exp \left(-i \int d^4x \mathcal{L} \right), \quad (3)$$

where X represents the fields implicit in Γ^μ and G . These fields will from now on be suppressed in our notation.

A. A brief overview of the heavy-fermion method

If the fermion is heavy, we can for low four-momenta treat it as essentially static and expand around its large mass. Following Ref. [2], this is accomplished by writing the fermion momentum p^μ in terms of the four-velocity v^μ and a residual fermion momentum l^μ as

$$p^\mu = mv^\mu + l^\mu, \quad (4)$$

where $l^\mu \ll m$ and $v^2 = 1$. Consequently, for an on-shell fermion, $p^2 = m^2$ implies that

$$2mv \cdot l + l^2 = 0. \quad (5)$$

The fermion field is next split into large (H) and small (h) components

$$\psi = e^{-imv \cdot x}(H + h), \quad (6)$$

where $\not{v}H = H$ and $\not{v}h = -h$, which shifts all the linear mass dependence to the small component. Without loss of generality we will assume for the remainder of this article that $v^\mu = (1, 0, 0, 0)$, and thus $\not{v} = \gamma_0$, which simplifies the algebra and the comparison with FW. (It is possible to keep v^μ general and also do the FW transformation in terms of v^μ [10, 13], but we choose a simpler approach here.) The resulting Lagrangian is

$$\mathcal{L} = \begin{pmatrix} H^\dagger & h^\dagger \end{pmatrix} \begin{pmatrix} i\mathcal{D}_0 + A & -i\vec{\sigma} \cdot \vec{\mathcal{D}} + B \\ -i\vec{\sigma} \cdot \vec{\mathcal{D}} + C & 2m + i\mathcal{D}_0 + D \end{pmatrix} \begin{pmatrix} H \\ h \end{pmatrix}, \quad (7)$$

where we have chosen to work with H^\dagger instead of \bar{H} to ease the comparison with the FW expressions further on. The cross talk between upper and lower components can be eliminated by defining a new small component field h' [3] such that

$$\begin{aligned} h &= h' - (2m + i\mathcal{D}_0 + D)^{-1}(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)H, \\ h^\dagger &= h'^\dagger - H^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(2m + i\mathcal{D}_0 + D)^{-1}. \end{aligned} \quad (8)$$

This turns the Lagrangian into a block-diagonal form:

$$\mathcal{L} = H^\dagger \left(i\mathcal{D}_0 + A - (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B) \frac{1}{2m + i\mathcal{D}_0 + D} (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C) \right) H + h'^\dagger (2m + i\mathcal{D}_0 + D) h', \quad (9)$$

and ‘completes the square’ in the path integral exponential. Up to this point no approximations have been made, so this equation can be regarded as a Dirac equation in two-component

form. After ‘integrating out’ the small components, which gives a constant determinant [3] (see the Appendix), only the H fields remain and we have obtained a rudimentary form of the HF Lagrangian. Note that the H fields have to be renormalized to preserve the norm [14, 21]. The antifermion is implicitly included as the ‘ z -graph’-like term of Eq. (9). We next expand this term, assuming that $i\mathcal{D}^\mu, G \ll 2m$. Thus, this part of the HF Lagrangian is expanded in powers of $(i\mathcal{D}_0 + D)/2m$, which contains only time derivatives of the fermion field (apart from possible derivatives in G), sandwiched between two space derivatives. The first few orders of the HF Lagrangian are then

$$\mathcal{L}_{\text{HF}}^{(0)} = H^\dagger(i\mathcal{D}_0 + A)H, \quad (10)$$

$$\mathcal{L}_{\text{HF}}^{(1)} = -\frac{1}{2m}H^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)H, \quad (11)$$

$$\mathcal{L}_{\text{HF}}^{(2)} = \frac{1}{4m^2}H^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(i\mathcal{D}_0 + D)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)H, \quad (12)$$

where the indices indicate inverse powers of m . The Lagrangian in general also contains higher order interaction terms in A - D and \mathcal{D}^μ . This expansion will not be shown explicitly. For the purposes of this paper we count $\mathcal{D}_0 \sim l_0 \sim |\mathbf{l}|$, although counting $\mathcal{D}_0 \sim \mathbf{l}^2/m$ is also possible, depending on the process being studied. Short-range interactions are parametrized as local counter terms, which contain LECs. The LECs are determined either by matching conditions (when the underlying, more complete theory is well known, e.g., HQEFT and non-relativistic QED (NRQED) [22]) or by fitting EFT calculations to data (when the underlying theory is not well known, e.g., HB χ PT). In some cases, models such as resonance saturation are used to evaluate or estimate the LECs.

The propagator for the H field is given by Eq. (9)

$$\frac{i}{i\partial_0 + \nabla^2 \frac{1}{2m + i\partial_0}} = \frac{i(1 + \frac{i\partial_0}{2m})}{i\partial_0 - \frac{\square}{2m}}, \quad (13)$$

where $\square = \partial_0^2 - \nabla^2$. The inverse of this propagator is non-linear in $i\partial_0$ beyond a certain order, and this propagator usually appears in expanded form [7, 19, 20]. In order to compare this HF propagator to the FW propagator that we will give below, we choose a situation that appears often in our own line of research, i.e., pion-nucleon dynamics in HB χ PT. Consider an on-shell nucleon with residual four-momentum l . The HF nucleon propagator after it has emitted one pion of four-momentum q (real or virtual, but with $q_0, |\mathbf{q}| \ll m$) can be written

as

$$\frac{i \left(1 + \frac{l_0 - q_0}{2m}\right)}{l_0 - q_0 + \frac{(l - q)^2}{2m}} = \frac{i \left(1 + \frac{l_0 - q_0}{2m}\right)}{-q_0 - \frac{l \cdot q}{m} + \frac{q^2}{2m}}, \quad (14)$$

where the relativistic on-shell condition (5) has been applied. The propagator (14) can be further expanded as needed, whether in $1/m$ or in $\sqrt{m_\pi/m}$ as in pion production off two nucleons [23]. This application of the propagator is similar to the one advocated by Hanhart and Wirzba [24]. After a simple re-identification of l and q as the electron and photon four-momenta, Eq. (14) can also represent the NRQED electron propagator after emission of a soft photon.

The HF propagator [Eq. (14)] is identical to an $1/m$ expansion of the relativistic propagator, projected on the upper left 2×2 submatrix:

$$\begin{aligned} \frac{i(m\not{v} + \not{l} - \not{q} + m)}{(mv + l - q)^2 - m^2} &\rightarrow i \frac{2m + l_0 - q_0}{-2(m + l_0)q_0 + 2\mathbf{l} \cdot \mathbf{q} + q^2} \\ &= i \frac{1 + \frac{l_0 - q_0}{2m}}{-q_0 - \frac{l \cdot q}{m} + \frac{q^2}{2m}}. \end{aligned} \quad (15)$$

B. Foldy-Wouthuysen transformation in a Lagrangian formulation

In order to compare to the HF results we choose to derive the FW transformation from the Lagrangian [10, 15] instead of the Hamiltonian as is usually done. In addition this approach makes it possible to treat the time and space components in a more symmetrical manner. The time derivative is implicit in all expressions containing \mathcal{D}_0 . Our starting point is the Lagrangian (1) rewritten as

$$\mathcal{L} = \psi^\dagger \left(i\mathcal{D}_0 - i\gamma_0 \vec{\gamma} \cdot \vec{\mathcal{D}} - \gamma_0 m + G \right) \psi, \quad (16)$$

where the field ψ^\dagger is chosen instead of $\bar{\psi}$ since it makes the algebra somewhat more transparent for the present purposes. This expression can be further separated into odd and even terms, i.e., terms that mix and do not mix the upper and lower components (in the Dirac representation)

$$\mathcal{L} = \psi^\dagger (-\mathcal{E} - \mathcal{O} - \gamma_0 m) \psi, \quad (17)$$

where the sign convention for the even (\mathcal{E}) and odd (\mathcal{O}) operators is in accordance with Bjorken-Drell [9]. The mass term can be removed from the upper components by shifting

the zero-point of the energy, i.e., we define $\tilde{\psi} = e^{-imt}\psi$, giving

$$\mathcal{L} = \tilde{\psi}^\dagger [-\mathcal{E} - \mathcal{O} - (\gamma_0 - 1)m] \tilde{\psi}. \quad (18)$$

The explicit expressions for the even and odd operators are

$$\mathcal{E} \equiv - \begin{pmatrix} i\mathcal{D}_0 + A & 0 \\ 0 & i\mathcal{D}_0 + D \end{pmatrix}, \quad (19)$$

$$\mathcal{O} \equiv - \begin{pmatrix} 0 & -i\vec{\sigma} \cdot \vec{\mathcal{D}} + B \\ -i\vec{\sigma} \cdot \vec{\mathcal{D}} + C & 0 \end{pmatrix}. \quad (20)$$

The FW transformation redefines the fermion field through the consecutive application of unitary matrices:

$$\psi''' = e^{iS''} e^{iS'} e^{iS} \tilde{\psi}, \quad (21)$$

$$S = -\frac{i\gamma_0 \mathcal{O}}{2m}, \quad (22)$$

where S' is related to \mathcal{O}' in the same way as S to \mathcal{O} [Eq. (22)] and similarly for S'' and \mathcal{O}'' . These higher order odd operators are introduced below. The transformations carried out to $\mathcal{O}(1/m^3)$ lead to the Lagrangian [9]

$$\mathcal{L}''' = \psi'''^\dagger (-\mathcal{E}' - (\gamma_0 - 1)m) \psi''', \quad (23)$$

$$\mathcal{E}' = \gamma_0 \left(\frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3} \right) + \mathcal{E} - \frac{1}{8m^2} [\mathcal{O}, [\mathcal{O}, \mathcal{E}]], \quad (24)$$

$$\mathcal{O}' = \frac{\gamma_0}{2m} [\mathcal{O}, \mathcal{E}] - \frac{\mathcal{O}^3}{3m^2}, \quad (25)$$

$$\mathcal{O}'' = \frac{\gamma_0}{2m} [\mathcal{O}', \mathcal{E}'], \quad (26)$$

where it is assumed that the time derivative in \mathcal{E} , i.e., in \mathcal{D}_0 , acts on everything to its right. All odd terms have been transformed away to this order and only the even operators remain. The explicit expressions for the higher order odd operators are

$$\mathcal{O}' = \frac{1}{2m} \begin{pmatrix} 0 & B' \\ -C' & 0 \end{pmatrix}, \quad (27)$$

$$\mathcal{O}'' = -\frac{1}{4m^2} \begin{pmatrix} 0 & B'(i\mathcal{D}_0 + D) - (i\mathcal{D}_0 + A)B' \\ C'(i\mathcal{D}_0 + A) - (i\mathcal{D}_0 + D)C' & 0 \end{pmatrix}, \quad (28)$$

where we have defined

$$B' = (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(i\mathcal{D}_0 + D) - (i\mathcal{D}_0 + A)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B), \quad (29)$$

$$C' = (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(i\mathcal{D}_0 + A) - (i\mathcal{D}_0 + D)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C). \quad (30)$$

The upper (N) and lower (n) FW components of ψ''' [Eq. (21)] are related to the HF fields H and h by the transformation

$$\begin{pmatrix} N \\ n \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{8m^2}\tilde{B}\tilde{C} + \frac{1}{8m^3}B'\tilde{C} & -\frac{1}{2m}\tilde{B} + \frac{1}{4m^2}B' + \frac{3}{16m^3}\tilde{B}\tilde{C}\tilde{B} - \frac{1}{8m^3}B'' \\ \frac{1}{2m}\tilde{C} + \frac{1}{4m^2}C' - \frac{3}{16m^3}\tilde{C}\tilde{B}\tilde{C} + \frac{1}{8m^3}C'' & 1 - \frac{1}{8m^2}\tilde{C}\tilde{B} - \frac{1}{8m^3}C'\tilde{B} \end{pmatrix} \begin{pmatrix} H \\ h \end{pmatrix}, \quad (31)$$

where we use the abbreviated notation

$$\tilde{B} \equiv -i\vec{\sigma} \cdot \vec{\mathcal{D}} + B, \quad (32)$$

$$\tilde{C} \equiv -i\vec{\sigma} \cdot \vec{\mathcal{D}} + C, \quad (33)$$

$$B'' \equiv B'(i\mathcal{D}_0 + D) - (i\mathcal{D}_0 + A)B', \quad (34)$$

$$C'' \equiv C'(i\mathcal{D}_0 + A) - (i\mathcal{D}_0 + D)C'. \quad (35)$$

The Lagrangian after the three FW transformations is

$$\mathcal{L}_{\text{FW}}^{(-1)} = 2mn^\dagger n, \quad (36)$$

$$\mathcal{L}_{\text{FW}}^{(0)} = N^\dagger(i\mathcal{D}_0 + A)N + n^\dagger(i\mathcal{D}_0 + D)n, \quad (37)$$

$$\begin{aligned} \mathcal{L}_{\text{FW}}^{(1)} &= -\frac{1}{2m}N^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)N \\ &\quad + \frac{1}{2m}n^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)n, \end{aligned} \quad (38)$$

$$\begin{aligned} \mathcal{L}_{\text{FW}}^{(2)} &= \frac{1}{4m^2}N^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(i\mathcal{D}_0 + D)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)N \\ &\quad - \frac{1}{8m^2}N^\dagger\{(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C), (i\mathcal{D}_0 + A)\}N \\ &\quad + \frac{1}{4m^2}n^\dagger(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(i\mathcal{D}_0 + A)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B)n \\ &\quad - \frac{1}{8m^2}n^\dagger\{(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B), (i\mathcal{D}_0 + D)\}n. \end{aligned} \quad (39)$$

By construction terms mixing upper and lower components start to appear only at the next higher order, so the square is already ‘completed’ and one can integrate out the $n^\dagger n$ contributions to this order. As in the HF case, the integration over the small components will

give a constant determinant (see the Appendix). When higher order terms, involving higher order LECs, are included in the relativistic Lagrangian, one has to take into account that A - D and \mathcal{D}^μ contain higher order interaction terms, such that further expansion of these quantities becomes necessary. Just like HF, the FW transformation generates an expansion in $1/m$, but the two series are not identical.

For example, the free N propagator is (see Ref. [9])

$$\frac{i}{i\partial_0 - (\sqrt{m^2 - \nabla^2} - m)} = \frac{i}{i\partial_0 + \frac{\nabla^2}{2m} + \frac{\nabla^4}{8m^3} + \dots}, \quad (40)$$

i.e., exactly what would naively be expected for a non-relativistic propagator. Unlike the inverse of the HF propagator (13) the inverse of (40) is linear in $i\partial_0$, although its dependence on the spatial components is more complicated. Applying this FW propagator to the same situation as in the HF case (but with the relativistic on-shell condition now being $l_0 = \sqrt{m^2 + \mathbf{l}^2} - m = \frac{\mathbf{l}^2}{2m} - \frac{\mathbf{l}^4}{8m} \dots$) we get

$$\frac{i}{l_0 - q_0 - \frac{(\mathbf{l}-\mathbf{q})^2}{2m} + \frac{(\mathbf{l}-\mathbf{q})^4}{8m^3} + \dots} = \frac{i}{-q_0 - \frac{\mathbf{q}^2 - 2\mathbf{l}\cdot\mathbf{q}}{2m} + \frac{(\mathbf{l}-\mathbf{q})^4 - \mathbf{l}^4}{8m^3} + \dots}, \quad (41)$$

which differs from the HF propagator [Eq. (14)]. Although the HF and FW propagators go on-shell for the same value of l^μ for any given order, in general they do differ off-shell beyond leading order. Also the on-shell residues are different, as discussed already in [14]. It was pointed out in Ref. [14], however, that the difference between these propagators can be compensated for by a field redefinition.

C. Comparison of HF and FW to third order

In contrast to HF, FW is an expansion in terms of both space and time derivatives of the fermion, i.e., the FW transformation is not restricted to the time derivatives only. The two schemes also differ in that the integrated-out pieces of FW have contributions at all orders, not just the first few as in HF—compare Eq. (9) to Eqs. (36)–(39). Comparison with Eq. (12) shows that the difference in the Lagrangian between these two expansion choices is given (to this order) by the second line of Eq. (39), i.e., differences start to appear at $\mathcal{O}(1/m^2)$. Both result in expansions in $1/m$, where m indicates the size of the cutoff of the theory. Note that the fields H and N are not identical (and neither are h' and n), except at the lowest order, but are related through a rather involved transformation, see

below. In contrast to the transformation in [14], which was carried out to $\mathcal{O}(1/m^5)$ for a constant electric field, we have kept the Lagrangian general (not necessarily renormalizable, e.g., HB χ PT), but restricted the expansion to $\mathcal{O}(1/m^2)$.

In order to transform from the HF to FW formalisms we need to ‘complete the square’ for the H and h fields according to Eq. (8). This boils down to an additional transformation

$$\begin{pmatrix} H \\ h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{2m+i\mathcal{D}_0+D}(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C) & 1 \end{pmatrix} \begin{pmatrix} H \\ h' \end{pmatrix}. \quad (42)$$

We then arrive at the field transformation

$$\psi^{(i)} = \begin{pmatrix} N^{(i)} \\ n^{(i)} \end{pmatrix} = M^{(i)} \begin{pmatrix} H \\ h' \end{pmatrix}, \quad (43)$$

where

$$M^{(1)} = \begin{pmatrix} 1 & -\frac{1}{2m}\tilde{B} \\ 0 & 1 \end{pmatrix}, \quad (44)$$

$$M^{(2)} = \begin{pmatrix} 1 + \frac{1}{8m^2}\tilde{B}\tilde{C} & -\frac{1}{2m}\tilde{B} + \frac{1}{4m^2}B' \\ \frac{1}{4m^2}\tilde{C}(i\mathcal{D}_0 + A) & 1 - \frac{1}{8m^2}\tilde{C}\tilde{B} \end{pmatrix}, \quad (45)$$

$$M^{(3)} - M^{(2)} = \begin{pmatrix} -\frac{1}{8m^3}\tilde{B}(i\mathcal{D}_0 + D)\tilde{C} & \frac{3}{16m^3}\tilde{B}\tilde{C}\tilde{B} - \frac{1}{8m^3}B'' \\ -\frac{1}{8m^3}\tilde{C}\tilde{B}\tilde{C} + \frac{1}{8m^3}[C' - (i\mathcal{D}_0 + D)\tilde{C}](i\mathcal{D}_0 + A) & -\frac{1}{8m^3}C'\tilde{B} \end{pmatrix}. \quad (46)$$

We have checked that this transformation indeed transforms the Lagrangian from the FW to the HF representation [to $\mathcal{O}(1/m^2)$]. This is *not* a unitary transformation, because of the block-diagonalizing done in Eq. (42). It is equivalent to the wave function renormalization and unitary field redefinitions in Refs. [14, 21]. In contrast to Ref. [14], however, we give the general transformation to $\mathcal{O}(1/m^2)$, i.e., it applies equally well to non-Abelian and non-renormalizable theories. Obviously, the FW spinor N contains parts of the (to be) integrated-out HF spinor h' . This means that transforming the upper/large components from one to the other mixes in lower components of the other representation. Thus, the transformation can be carried out explicitly only in cases where the lower component Lagrangian is known (up to the required order).

The general relation between the HF and FW Lagrangians to a given order is

$$H^\dagger U_{HH} H + h'^\dagger L_{h'h'} h' = N^\dagger U_{NN} N + n^\dagger L_{nn} n. \quad (47)$$

Assuming a general form for the i th order matrix $M^{(i)} = \begin{pmatrix} a^{(i)} & b^{(i)} \\ c^{(i)} & d^{(i)} \end{pmatrix}$ that relates the FW and HF fields [Eq. (43)] we can write [hereafter suppressing the index (i)]

$$\begin{pmatrix} U_{HH} & 0 \\ 0 & L_{h'h'} \end{pmatrix} = \begin{pmatrix} a^\dagger & c^\dagger \\ b^\dagger & d^\dagger \end{pmatrix} \begin{pmatrix} U_{NN} & 0 \\ 0 & L_{nn} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (48)$$

This puts constraints on the Lagrangian operators U_{XX} and L_{xx} :

$$U_{HH} = a^\dagger U_{NN} a + c^\dagger L_{nn} c, \quad (49)$$

$$L_{h'h'} = b^\dagger U_{NN} b + d^\dagger L_{nn} d, \quad (50)$$

$$0 = a^\dagger U_{NN} b + c^\dagger L_{nn} d, \quad (51)$$

$$0 = b^\dagger U_{NN} a + d^\dagger L_{nn} c. \quad (52)$$

By taking advantage of the block-diagonal conditions [(51) and (52)], the lower component L_{nn} can be eliminated from the expression for U_{HH} and we arrive at a relation between the upper left blocks of the two Lagrangians:

$$U_{HH} = a^\dagger U_{NN} (a - b d^{-1} c) = [a^\dagger - c^\dagger (d^\dagger)^{-1} b^\dagger] U_{NN} a. \quad (53)$$

This gives a relation between the operator forms of the Lagrangians that remain after integrating out the respective lower components. To $\mathcal{O}(1/m^2)$ this expression is simplified to

$$U_{HH} = a^\dagger U_{NN} a. \quad (54)$$

The explicit expression as provided by Eq. (45) is

$$U_{HH} = \left(1 + \frac{1}{8m^2} \tilde{B} \tilde{C}\right) U_{NN} \left(1 + \frac{1}{8m^2} \tilde{B} \tilde{C}\right), \quad (55)$$

which is the general form of the wave function renormalization to this order. It reproduces, e.g., the special cases considered in [14]. Incidentally, this expression explains why the $\mathcal{O}(1/m)$ expressions look equivalent—the transformation between the HF and FW Lagrangians is $\mathcal{O}(1/m^2)$.

The $1/m$ corrections to contact terms can be different for the two methods. This can be seen in, e.g., the πN Lagrangian that has been developed to $\mathcal{O}(1/m^3)$ [19, 25] and higher [20] in HB χ PT, using the HF method. We find that the g_A^3/m^2 corrections to the LECs \hat{d}_{12} and

\hat{d}_{13} in [19, 20] are different in the HF and FW formalisms. Thus, if these LECs are calculated, e.g., by resonance saturation, or extracted from data using the relativistic formulation, it would seem like the corresponding $1/m$ corrections would matter. This difference could be absorbed by a field renormalization and unitary transformation, see Refs [14, 21], provided the lower components of the relativistic Lagrangian are known. Therefore, the equivalence of the HF and FW approaches can be established quite generally and in particular regarding the values of LECs at higher orders. Since the FW result reproduces the relativistic S -matrix [17], it is clear that so will the corresponding HF result.

III. CONCLUSIONS

We have carried out the heavy-fermion (HF) and Foldy-Wouthuysen (FW) non-relativistic reduction schemes to $\mathcal{O}(1/m^2)$, starting from a quite general, relativistic, Lagrangian. At this order explicit differences start to appear in the expressions of the respective non-relativistic Lagrangians. Our investigation treats a more general Lagrangian than the one in [14] and is especially concerned with the LECs (introduced in the relativistic Lagrangian) as they appear in the NR theory, e.g., χ PT.

We have derived the field transformation between the two descriptions to $\mathcal{O}(1/m^2)$ and have shown how the $1/m$ corrections to certain LECs differ between the two approaches at this order. Although the Lagrangians differ at $\mathcal{O}(1/m^2)$, they can be related through a field redefinition, which we give in a more general form than the one in Ref. [14]. Thus, when using LECs of higher order, one must ascertain that the fields are appropriately redefined.

We have also shown the equivalence between the NR propagators derived in HF and from an $1/m$ expansion. On the other hand, even though the HF and FW propagators go on-shell for the same four-momentum, they will differ off-shell. Also this difference is captured by the field redefinition. In order to do the field redefinition it is necessary to know the lower components of the original, relativistic Lagrangian.

Acknowledgments

We appreciate discussions with Daniel Phillips and Harold Fearing. This work was supported in part by the NSF grant No. PHY-0457014.

APPENDIX: INTEGRATING OUT SMALL COMPONENTS

We repeat here the arguments given in Ref. [3] (see also [14]) showing that the integration over the lower components of the HF Lagrangian gives a constant. The discussion automatically carries over to the FW case. The path integral we need to perform for HF is

$$\begin{aligned}
\mathcal{I}_{\text{HF}} &= \int [dh'^{\dagger}][dh'] \exp \left[i \int d^4x h'^{\dagger} (2m + i\mathcal{D}_0 - i\epsilon) h' \right] \\
&= \text{Det}(2m + i\partial_0 + i\Gamma_0 - i\epsilon) = \text{Det}(2m + i\partial_0) \text{Det} \left(1 + \frac{1}{2m + i\partial_0 - i\epsilon} i\Gamma_0 \right) \\
&= \text{Det}(2m + i\partial_0) \exp \left[\frac{1}{2} \text{Tr} \ln \left(1 + \frac{1}{2m + i\partial_0 - i\epsilon} i\Gamma_0 \right) \right] = \text{Const.} \tag{A.1}
\end{aligned}$$

As argued in Ref. [3], the propagator in the logarithm is always propagating in the backward time direction only. Thus the trace can not be closed, it has to vanish, and the determinant is a constant as indicated.

To $\mathcal{O}(1/m^2)$ in FW we have

$$\begin{aligned}
\mathcal{I}_{\text{FW}} &= \int [dn^{\dagger}][dn] \exp \left[i \int d^4x n^{\dagger} \left(2m + i\mathcal{D}_0 + A + \frac{1}{2m} (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B) \right. \right. \\
&\quad + \frac{1}{4m^2} (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(i\mathcal{D}_0 + A)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B) \\
&\quad \left. \left. - \frac{1}{8m^2} \{ (-i\vec{\sigma} \cdot \vec{\mathcal{D}} + C)(-i\vec{\sigma} \cdot \vec{\mathcal{D}} + B), (i\mathcal{D}_0 + D) \} - i\epsilon \right) n \right]. \tag{A.2}
\end{aligned}$$

The resulting determinant can, just as in the HF case, be factorized into the free propagator and interaction terms:

$$\begin{aligned}
\mathcal{I}_{\text{FW}} &= \text{Det} \left(2m + i\partial_0 - \frac{\nabla^2}{2m} - \frac{\nabla^4}{8m^3} + \dots \right) \\
&\quad \times \text{Det} \left[1 + \frac{1}{2m + i\partial_0 - \frac{\nabla^2}{2m} - \frac{\nabla^4}{8m^3} + \dots - i\epsilon} (\text{interaction terms}) \right] \\
&= \text{Det} \left(2m + i\partial_0 - \frac{\nabla^2}{2m} - \frac{\nabla^4}{8m^3} + \dots \right) \\
&\quad \times \exp \left[\frac{1}{2} \text{Tr} \ln \left(1 + \frac{1}{2m + i\partial_0 - \frac{\nabla^2}{2m} - \frac{\nabla^4}{8m^3} + \dots - i\epsilon} (\text{interaction terms}) \right) \right] = \text{Const.}, \tag{A.3}
\end{aligned}$$

where in the last step we again use that the propagator propagates in the backward time

direction only, so that no closed loops are possible.

-
- [1] S. Weinberg, *Physica* **96A**, 327 (1979).
 - [2] H. Georgi, *Phys. Lett.* **B240**, 447 (1990).
 - [3] T. Mannel, W. Roberts, and Z. Ryzak, *Nucl. Phys.* **B368**, 264 (1992).
 - [4] N. Brambilla, A. Pineda, J. Soto, and A. Vairo, *Rev. Mod. Phys.* **77**, 1423 (2005).
 - [5] E. Jenkins and A. Manohar, *Phys. Lett.* **B255**, 558 (1991).
 - [6] V. Bernard, N. Kaiser, J. Kambor, and U.-G. Meißner, *Nucl. Phys.* **B388**, 315 (1992).
 - [7] V. Bernard, N. Kaiser, and U.-G. Meißner, *Int. J. Mod. Phys. E* **4**, 193 (1995).
 - [8] L. L. Foldy and S. A. Wouthuysen, *Phys. Rev.* **78**, 29 (1950).
 - [9] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill, (New York, 1964).
 - [10] J. G. Körner and G. Thompson, *Phys. Lett.* **B264**, 185 (1991).
 - [11] T. Mannel and G. A. Schuler, *Phys. Lett.* **B349**, 181 (1995); *Z. Phys.* **C67**, 159 (1995).
 - [12] S. Balk, J. G. Körner, and D. Pirjol, *Nucl. Phys.* **B428**, 499 (1994); S. Balk, A. Ilakovac, J. G. Körner, and D. Pirjol, [arXiv:hep-ph/9307230].
 - [13] A. Das and V. S. Mathur, *Phys. Rev. D* **49**, 2508 (1994).
 - [14] A. Das, *Mod. Phys. Lett. A* **9**, 341 (1994); [arXiv:hep-ph/9310372].
 - [15] B. R. Holstein, *Am. J. Phys.* **65**, 519 (1997).
 - [16] N. Brambilla, E. Mereghetti, and A. Vairo, *JHEP* **08**, 039 (2006).
 - [17] H. W. Fearing, G. I. Poulis and S. Scherer, *Nucl. Phys.* **A570**, 657 (1994); S. Scherer, G. I. Poulis and H. W. Fearing, *Nucl. Phys.* **A570**, 686 (1994).
 - [18] S. Ando, T.-S. Park, K. Kubodera, and F. Myhrer, *Phys. Lett.* **B533**, 25 (2002).
 - [19] N. Fettes, U.-G. Meißner, and S. Steininger, *Nucl. Phys. A* **640**, 199 (1998).
 - [20] N. Fettes, U.-G. Meißner, M. Mojžiš, and S. Steininger, *Ann. Phys.* **283**, 273 (2000); *ibid* **288**, 249(E) (2001).
 - [21] S. Okubo, *Prog. Theor. Phys.* **12**, 102 (1954); **12** 603 (1954).
 - [22] W. E. Caswell and G. P. Lepage, *Phys. Lett.* **B167**, 437 (1986).
 - [23] T. D. Cohen, J. L. Friar, G. A. Miller, and U. van Kolck, *Phys. Rev. C* **53**, 2661 (1996).
 - [24] C. Hanhart and A. Wirzba, arXiv:nucl-th/0703012.

[25] G. Ecker and M. Mojžiš, Phys. Lett. B **365**, 312 (1996).